實驗資料分析與電腦

沈青嵩　國立臺灣師範大學

一、緒　論

人類文化發展到一定程度後，對自已所處的環境總希望能徹底了解，對周遭的事事物物，期望能有合理的解釋，放眼觀察宇宙現象似乎錯綜複雜，漫無頭緒；仰頭觀望蒼穹之星辰運轉，詳細觀測周遭的各種聲、光、電、熱等現象，但見四季更迭井然有序，日月運轉周而復始，混亂中似乎蘊藏著某種次序，不禁使人們確信大自然各現象間隱存著某種因果關係。對自然界的現象加以深入的觀察、研究，然後演繹歸納出他們之間的因果關係，使我們澈底的了解宇宙，一則滿足人類的好奇心，進而可造福人群，提昇文明，這是科學家一致的目標。

一部物理學的發展史可說是實驗的發展史，刻爾文爵士（ Lord Kelvin ）曾說過“無準確的度量，即無科學”，回顧四百多年前哥白尼首倡地球轉動說，在各界紛加指責聲中，泰勒勃拉（ Tycho Brahe ）堅信：一個正確的模式或理論之成熟有賴於精確實驗數據之獲得，哥白尼理論正確與否唯有從事實中去獲得答案；於是他製造大型六分儀和羅盤針，並建立了當代世界最大的天文台，窮二十年之精力觀測行星，他的助手刻卜勒分析觀測所得之數據並計算水星之軌跡，刻卜勒根據哥白尼的學說地球做圓軌道公轉並加上地球自轉之修正，發現理論結果與實際數據有八分（ 1度60分 ）之誤差，但刻卜勒深信泰戈勃拉的觀測數據誤差不會超過二分，就由此六分的相差，使這位精於數學的科學家提出行星軌道爲橢圓而非正圓，太陽且在橢圓焦點之一等三大定律。一百年後牛頓也因刻卜勒定律而提出影響深遠的萬有引力定律以解釋天體運動，而使物理學在十七世紀中有燦爛的成果。十九世紀末更因量度的進步而發現古典力學在與光速可比較的高速下及小至100埃內的小範圍下有了很大的不準度，古典電磁學在高頻率及小凡度下也有了矛盾現象，而促使相對論與量子論的誕生，並開近代物理之先河。更精確的天文觀測，發現水星近日點每一世紀約移動四十三秒，這使愛恩斯坦的廣義相對論有了證據，由此可知精確量度與實驗數據處理之重要性。

很不幸目今中學實驗教學，非常的陌視實驗數據之處理，動手實驗由於各方面的提倡及大學入學考試必考的考試領導教學下已獲得中學教師與學生的重視，唯學生實驗大部份爲印證已知之定律，把所獲之數據充其量求一下平均值及百分誤差而已，很少進一步去分析，譬如就最簡單的線性函數作圖，我們就以印證彈簧恢復力與伸長量X成正比的虎克定律作例子，理想下F＝KX兩者成線性，故F－X圖上爲一直線，但在實驗上，由於儀器刻度或實驗者之偏差等產生的系統誤差（ Systematic errors ），數據無法個個均落在理想直線上，縱使系統誤差可完全校正與避免，而另一種我們不知其因的混亂誤差（random error ），在統計變動（ statistical flutuation ）之下，使同一人操作的實驗，重複再量度一次

，兩組實驗數據無法完全相同。假設實驗數據如圖 1 所示，如何從這些點中畫出一條最恰當的直線呢？一般學生可能直R拿來，憑自已的直覺信手繪出一條直線來，量量斜率那就是m值了，照

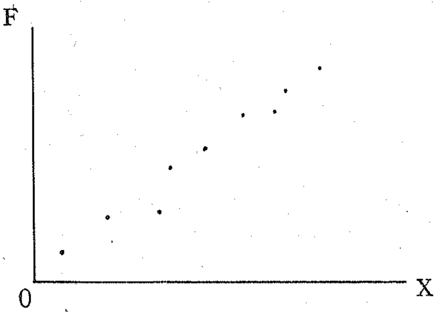

圖1：虎克定律

上述的直覺法，同一組數據，五十人來畫可能會畫出五十種斜率不同的直線來，所謂「失之毫釐，差之千里」，近代科學講究的是精確的量度，二十世紀物理的成就，大部份得力於量度儀器精密度的改進，譬如由於分光儀的改進，同一條明線光譜經分析後可見它分成數條而發現了Fine Structure 及 hyperfine structure ，而引進了電子自轉及原子核自轉的電磁效應使量子力學更上一層樓。再說科學的態度應該是客觀的，有根據的畫出最恰當的直線，那就需要用最小二次式固定法（Least-squares Fit ）來找出該直線的斜率及截距，尤其曲線為然，這時往往需要借助於電腦，否則計算上眞是煩不勝煩，最小二次式固定法在實驗數據分析上應用極廣，本文除介紹誤差的一些理論外主要即在介紹此方法，及如何利用電腦處理它。首先來談談電腦的問題。

二、電腦在教學上的運用

電腦是一種極端錯綜複雜的電子裝備，由數以萬計以邏輯線路為主件之Chip組成的，可接受儲存和執行等之指令，這些指令應以它能接受的語言來編寫成程式輸入後，經電腦自動邏輯編譯，在極短時間內即可按指令輸出所需要的資料。以往輸出輸入常用打卡方式，固可收永久存放，方便再用之效，但也有浪費不便之弊，一般學生均較喜歡使用打字於機紙上的輸出輸入方式，至於教學上則可與閉路電視連接，輸出輸入均可顯示在螢光幕上，全班同學一目了然，非常方便，至於常用或重要的程式，照樣可永久儲存於電腦內，須用時一個指令就可把它叫出於工作儲存區（Working storagearea ），以備運用。

電腦主要的功用有下列數項：(1)可迅速從事算術運算，加減乘除，次方式及三角函數等任何繁複運算，用人工或許須經年累月，而它可在幾秒內算出，更可利用一些如do loop 一再重複運作的指令，處理人力所不能及之運算，端視程式之設計而定。(2)它可從事邏輯思考而且非常迅速，當然這種思考能力是使用人賦予它的，而非它能自動思考，在一連串指令下它能延伸人類思考的能力。(3)它可儲存極多的資料與數據，需要時可隨時用指令讓其輸出顯示於螢光幕上，它甚至可整理這些資料，如令雜亂無章的數據按其大小順序排列，也可繪出兩組變數間的函數圖形等，(4)電腦可做兩件或多件事實之邏輯比較。在實驗運用上，我們大都利用它的第一及第三項功能。尤其近代電腦不但能從事矩陣運算，連逆矩陣（inverse matrix）也可用一個指令令其算出，處理數據及理論運算，非常方便，最小二次方式固定法就是利用電腦在這方面的能力，尤其是多項式的固定利用電腦省時省力。

電腦還可做各種有意義的遊戲，如打棒球、高爾夫球等遊戲，非常生動有趣，小孩子尤其樂此不疲，今日電腦在美國已經相當普遍，中小學生也都有機會操作，筆者曾見一小學生利用週日在大學電腦中心玩此類遊戲，由於興緻太濃，一局一局的打下去，用了一大堆的機紙，使得輪值人員不得不請他改用螢光幕輸出，以節省用紙，設計遊戲程式只要深諳機械語言即可設計，使用者更簡單了，只要照它的指示（當然是程式設計人設計的）輸入就可看出結果，有時連大人也都玩得

津津有味。更可設計一些益智性的遊戲，如心算等，讓學生在遊戲中學習，效果益彰；也可把教材設計成 loop，讓學生自學，直至答案正確才能自動通過，否則會一再重複並提示學生可能錯在那裏，學生在極高興緻下思索、自學，實在是最有效最有前途的教具，尤其實驗資料處理，更缺不了它。所以美國 1960 年只大約有一千台電腦，1965 年增至 35000台，1970 年達 75000台目前更是普遍；我國也日漸重視，現在雖只有幾所大學及學術研究機關有，但鑑於電腦使用必日漸普遍，且今教育部委託師大科學教育中心正進行的「高級中學科學課程研究計劃」中，所擬科學課程模式A中即有高二及高三加開電腦原理等半年課程之選修課，而此模式獲至最多支持。畢竟有朝一日電腦將深入你我的生活中。

三、實驗資料分析

什麼叫資料分析呢？簡言之即做實驗觀察所得之物理現象或物理量的比對及解釋前的分析工作，因大部份這方面的工作都是針對物理量而言，故本文介紹範圍也以物理量的分析為主，至於為什麼要有資料分析呢？為了比對及解釋，必須先確定所測得物理量的眞實性或可靠性。

1.平均值與標準差

今欲量度某一物理量，多次測量之結果，獲有 n 個不同的 x_i值，每個 x_i值出現之次數為 f_i；表示這組數據的中央趨勢（Central tendency）在統計上常用的有下列三種：

(A)模值（mode）：出現頻率最高的值。

(B)中值（median）：數據按順序排列，頻率居中所對應之值。

(C)平均值（mean）：即以各數據乘以出現頻率之總和再除以頻率總數，以數學式表之為

$$\overline{x}=\sum_{i=1}^{n} f_i\, x_i \Big/ \sum_{i=1}^{n} f_i$$

當然以平均值最為實用。此外數據離散（data dispersion）程度的描述可讓人了解此量度數據的可信度及誤差範圍，通常中學實驗教學往往採用百分誤差法，但百分誤差須先知道欲量度物理量之眞值（true value），實際物理研究，根本不知眞值若干，以理論值代替眞值，往往變成倒因為果，多少物理定律本身就是實驗定律呢！所以嚴格的說百分誤差只是一種為教學方便的權宜方法，在統計上不被採用。描述數據離散程度在一般統計上有下列四種：

(A)範圍（range）：最高值與最低值之差。

(B)內半四分範圍（Semi - interquartile range）：先將數據按大小順序排列，根據頻率分成四等分，中間50％ 最高值與最低值差之稱。

(C)平均差（average deviation）：即

$$d=\left(\sum_{i=1}^{n} f_i\,|x_i-\overline{x}|\right)\Big/\sum_{i=1}^{n} f_i\ ,$$

(D)標準差（Standard deviation）：定義為

$$\sigma=\left(\sum_i f_i\,(x_i-\overline{x})^2\Big/\sum_i f_i\right)^{1/2}$$

各數據與平均值之差的平方再開方，可省卻平均差所規定的絕對值演化之麻煩，所以也是一般使用最廣的表示誤差範圍的方法，不過在取樣情況下分母應再減 1。即

$$S=\left[\sum_i f_i(x_i-\overline{x})^2\big/(N-1)\right]^{1/2},$$

稱為樣本標準差（Sample standard deviation），不過一般量度為求準確

$$N\left(=\sum_{i=1}^{n} f_i\right)$$

值應該相當大，分母減不減 1，結果相差無幾，只是在理論上有所別而已，所以以下就不再區別它們了。

在 N 極大的情況下，作 f 對 x 的分配圖，我們發現此分配圖形為高斯分配圖形也就是一般所說的常態分配，如圖 2 所示，越接近平均值 $\overline{x}$ 的數據出現的次數越多，換言之，數據靠近平均值

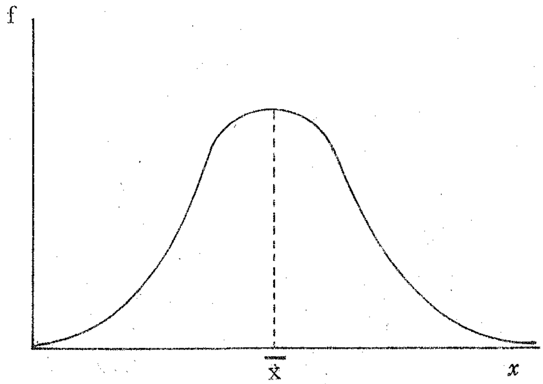

圖 2：高斯分配圖形

的或然率P較大，關於高斯分配（Gaussian distribution)容後詳述；為了讓他人了解你量度的結論，一般我們將其結果寫成 $\bar{x}\pm\sigma$ ，$\bar{x}$ 為平均值，σ 為標準差，σ 表示誤差的範圍；根據高斯分配函數計算，1σ 佔總分配的 34.13 %，如圖 3 所示；故 $\bar{x}\pm 1\sigma$ 等於給別人全部數據的 68.26 %的消息。其餘偏離平均值太遠的 31.74 %的數據不被採納。

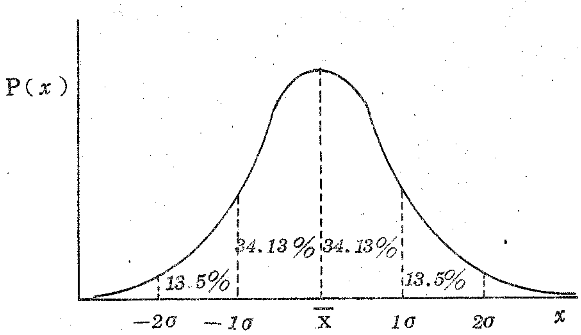

圖 3

現在考慮下列量度，對一物理量每次量有N個數據共量M次，則我們有M組數據，若M夠大，我們發現每組的平均值 $\bar{x}_\mu$ 對總數據的平均值 $\bar{x}$ 構成高斯分配關係，（在MN之值無限大時，$\bar{x}$ 可視為該物理量的眞值）這時的標準差稱為平均值標準差 σ_m（Standard deviation of the mean）或稱標準誤（Standard error），它很容易被證明 $\sigma_m=\frac{\sigma}{\sqrt{N}}$，今簡述於下：照定義

$$\sigma_m^2=\frac{1}{M}\sum_{\mu=1}^{M}(\bar{x}_\mu-\bar{x})^2=\frac{1}{M}\sum_{\mu=1}^{M}D_\mu^2$$

但

$$D_\mu=\bar{x}_\mu-\bar{x}=\left(\frac{1}{N}\sum_{i=1}^{N}x_{\mu i}\right)-\bar{x}=\frac{1}{N}\sum_{i=1}^{N}(x_{\mu i}-\bar{x})$$

$x_{\mu i}-\bar{x}$ 可視為各數據對總平均 $\bar{x}$ 的偏離，定為 $d_{\mu i}$

$$\therefore D_\mu^2=\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}d_{\mu i}d_{\mu j}=\frac{1}{N^2}\left(\sum_{i=1}^{N}d_{\mu i}^2+\sum_{i\neq j}^{N}d_{\mu i}d_{\mu j}\right)$$

在高斯分配中因各數據對 $\bar{x}$ 成對稱，故上式最後一項為零，所以標準誤可簡化為

$$\sigma_m^2=\frac{1}{MN^2}\sum_{\mu=1}^{M}\sum_{i=1}^{N}d_{\mu i}^2=\frac{\sigma^2}{N}$$

$$\therefore \sigma_m=\frac{\sigma}{\sqrt{N}}$$

標準誤的意義在說明一組數據平均值的可信程度，取樣平均值 $\bar{x}_\mu$（Sample means）與群數平均值 $\bar{x}$（population mean）相去的程度為何？我們可看出，量的數據越多，誤差越小；在T測試中（T Test）規定 t 值等於

$$\frac{\bar{x}_\mu-\bar{x}}{\sigma_m}$$

，這主要是應用在教育統計上，以顯示取樣的可信程度。

為何用一物理量多次量度得到的值無法完全相同呢？產生誤差的原因如上所述有系統誤差與混亂誤差，前者是因實驗者之姿勢等或儀器本身之偏差所引起的，可儘量修正，如讀R標時儘量垂

直對準欲讀之刻度，如因溫度因素所產生之熱膨脹。引起R標不準確可用公式校正；但後者則因統計上的誤差，無可避免，因此量度的結果往往是要標明誤差範圍的，如桌子長爲（15.3±0.8）cm，0.8厘米即爲標準差，15.3厘米即爲平均值。但是複雜的物理量往往不是直接可量出來的，而是許多可量量的函數，如動量爲質量與速度的乘積，而速度又爲距離與時間的函數，量出來的質量、距離與時間均有誤差，那該怎麼樣來表示這種物理量的誤差呢？這就得討論誤差的傳遞了。（Propagation of errors）。

2. 誤差的傳遞

假設一物理量Q是衆多可測量 a, b, c …之函數 $Q=f(a,b,c\cdots)$

則Q的變化度（Variance）照定義應爲

$$\sigma_Q^2=\frac{1}{N}\sum_{i=1}^{N}(Q_i-\overline{Q})^2$$ 式中 Q_i爲對應第 i 次測量量 a_i, b_i, c_i ……所算出之Q值，$\overline{Q}$爲N個 Q_i 的平均值，上式可化爲

$$\therefore \sigma_Q^2=\frac{1}{N}\sum_{i=1}^{N}\Delta Q_i^2=\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\partial Q}{\partial a}\Delta a_i+\frac{\partial Q}{\partial b}\Delta b_i+\cdots\cdots\right)^2$$

$$=\frac{1}{N}\sum_{i=1}^{N}\left\{\left[\left(\frac{\partial Q}{\partial a}\right)^2\Delta a_i^2+\left(\frac{\partial Q}{\partial b}\right)^2\Delta b_i^2+\cdots\cdots\right]+2\left(\frac{\partial Q}{\partial a}\frac{\partial Q}{\partial b}\Delta a_i\Delta b_i+\cdots\cdots\right)\right\}$$

式中 Δa_i 可視爲 $a_i-\overline{a}$，同理 $\Delta b_i=b_i-\overline{b}$，……假設量度次數極多，則 a_i，b_i 分別對 $\overline{a}$，$\overline{b}$ 成高斯分配，則因對稱緣故，上式右邊第二項爲零，故Q的變化度簡化爲

$$\sigma_Q^2=\left(\frac{\partial Q}{\partial a}\right)^2\sigma_a^2+\left(\frac{\partial Q}{\partial b}\right)^2\sigma_b^2+\left(\frac{\partial Q}{\partial C}\right)^2\sigma_c^2+\cdots\cdots\cdots\cdots$$

例如欲求距形方桌的周長P，測得數據整理後得：長平均 $\overline{L}=80$ cm 標準差 $\sigma_L=2$cm，寬平均 $\overline{W}=50$ cm，標準差 $\sigma_W=1$cm，則周長標準差爲

$$P=ZL+ZW$$

$$\therefore \sigma_P^2=\left(\frac{\partial P}{\partial L}\right)^2\sigma_L^2+\left(\frac{\partial P}{\partial W}\right)^2\sigma_W^2$$

$$=4\sigma_L^2+4\sigma_W^2=20\ \text{cm}^2$$

$$\therefore \sigma_P=4.5\ \text{cm}$$

$$\therefore P=(260\pm4.5)\ cm$$

3. 分配函數（Distribution function）

爲了更深入的了解，不得不簡要介紹數據出現次數（即頻率 frequency）與測量值間分佈的模式，以便與我們量得之數據比較歸類。

擲骰子 n 次中若有 x 次點數恰爲 1 的或然率爲 $\binom{n}{x}P^x q^{n-x}$ 式中 P爲每次成功的或然率，q 爲失敗率，本問題分別 $P=\frac{1}{6}$，$q=\frac{5}{6}$。事實上這就是統計上稱的二項式分配（binomial distribution）寫爲

$$P_B(x,n,p)=\binom{n}{x}p^x q^{n-x}$$

它的命名是因與二項式展開定律 $(p+q)^n=\sum_{x=0}^{n}\left[\binom{n}{x}p^x q^{n-x}\right]$

非常接近的緣故，數據成二項式分配時，其平均 u 應爲 np 即

$$u=\sum_{x=0}^{n}x\,P_B(x,n,p)=\sum_{x=0}^{\infty}\left[x\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}\right]=np$$

二項式分配若P極小時就成了波義遜分配（Poisson distribution）簡要說明如下：

$$P_B(x,n,p)=\frac{1}{x!}\frac{n!}{(n-x)!}P^x(1-P)^{-x}(1-P)^n$$

其中 $\frac{n!}{(n-x)!}=n(n-1)\cdots\cdots(n-x+1)$

共有 x 項相乘，當 $P\ll1$，就是討論 $x\ll n$的區間，故上項乘積可視爲 n^x，又

$$\lim_{p\to0}(1-p)^n=\lim_{p\to0}\left[(1-p)^{\frac{1}{p}}\right]^{\mu}=e^{-\mu}$$

$$\therefore \lim_{p\to 0} P_B(x,n,p)=\frac{(np)^x}{x!}e^{-\mu}=\frac{\mu^x e^{-\mu}}{x!}$$

$$=P_P(x)$$

由數學演化可證明，成波義遜分配數據的平均值 $\overline{x}$ 及標準差分別爲 $\overline{x}=\mu$ ，及 $\sigma=\sqrt{\mu}$ 圖 4 卽爲 $P_P(x)$ 之分配圖形

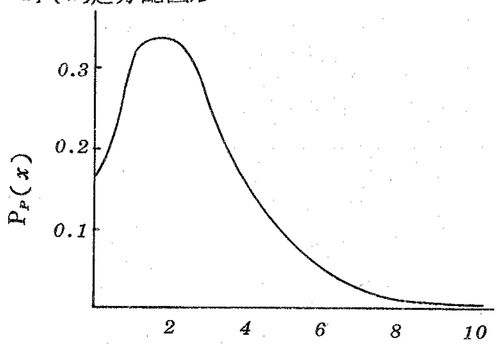

圖 4 波義遜分配函數： $\mu=1.67$， $\sigma=1.29$

放射性元素蛻變學習中，蛻變率測定的數據卽成波義遜分配，蓋 1 毫克的放射性元素有 10^{19} 原子，而在固定短時距蛻變的原子數與 10^{19} 相較是微乎其微。另一例子如核反應中，被加速粒子若每秒有 10^{13} 個射向靶（target），而可參與反應數目不到 10^4 個，故單位時間與靶反應的數目自然成波義遜分配了。

在數據統計分析上最重要的分配函數首推高斯分配，也就是一般所稱的常態分配（Normal distribution），前面的敍述也淸楚的告訴我們它的重要處。理論上當 n 趨近無限大而 p 保持相當之值而使 $np\gg 1$ 時，高斯分配就趨近於二項式分配，以下的導引可看出它的特性來，設高斯分配函數爲

$$P_g(x)=Ae^{-h^2(x-m)^2}$$

因總或然率爲 1 之故，所以

$$\int_{-\infty}^{\infty} P_g(x)dx=1 \Rightarrow A=\frac{h}{\sqrt{\pi}}$$

而平均值 $\overline{x}=\int_{-\infty}^{\infty} x\,P_g(x)dx \Rightarrow m=\overline{x}$

標準差 $\sigma^2=\int_{-\infty}^{\infty}(x-\overline{x})^2 P_g(x)dx \Rightarrow \sigma^2=\frac{1}{2h^2}$

因此 $P_g(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\overline{x})^2}{2\sigma^2}}$

這就是有名的高斯分配函數，明顯地，x 對 $\overline{x}$ 對稱，而 $f(x)$ 當 $x=\overline{x}$ 時有極大值，其轉點（inflection point）在 $x=\overline{x}\pm\sigma$ 處 其分佈圖形如圖 5 所示，高斯分配之重要性是一物理量在極多次量度後所成的分佈大都遵守此函數，教育統計應用更廣。

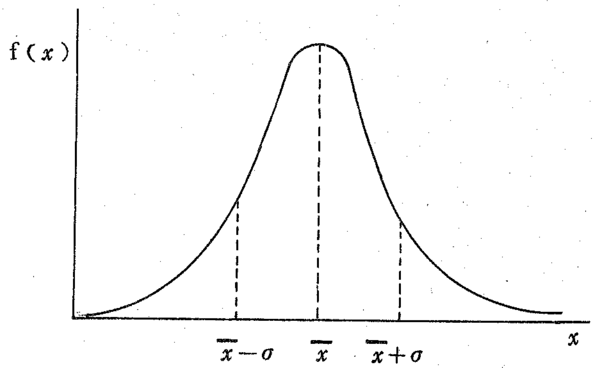

圖 5 ：高斯分配函數

4. 線性最小二次方式固定法（linear least-squares fit）

現在談到正題，設有一組數據爲

$$x_1, y_1 ; x_2, y_2 ; \cdots\cdots ; x_i, y_i ; \cdots\cdots ; x_n y_n$$

而我們相信 y 與 x 間存有 $y=f(x)$ 的函數關係，假設在量度上 y 的誤差大於 x 很多時，卽

$$\frac{\sigma_{yi}}{y_i} \gg \frac{\sigma_{xi}}{x_i}$$

，就只須考慮到 y 的標準差令其爲 σ，而忽略 x 的標準差。在 n 相當大的情況下也就是量度的次數相當多時， y_i 對 y 成高斯分配，則

$$P_i(y)=\frac{1}{\sigma_i\sqrt{2\pi}}\ exp\left\{-\tfrac{1}{2}\left(\frac{y_i-y}{\sigma_i}\right)^2\right\}$$

n 次量度下之或然率爲

$$P_t=P_1P_2\cdots\cdots P_n=\left(\pi_{i=1}^{n}\frac{1}{\sigma_i\sqrt{2\pi}}\right)$$

$$e\times p\ \{-\tfrac{1}{2}(\frac{y_i-y}{\sigma_i})^2\}$$

假設我們的量度是一組最接近 $y=f(x)$ 關係的量度，則 P_i 必須是極大值，此稱爲 Principle of maximum likelihood．欲 P_i 極大只須討論

$$D^2=\sum_{i=1}^{n}(\frac{y_i-y}{\sigma_i})^2$$

極小就可。這也就是稱爲最小二次方式固定法的原因。

在線性情況下　$y=f(x)=mx+c$　所以

$$D^2=\sum_{i=1}^{n}\frac{1}{\sigma_i^2}[y_i-(mx_i+c)]^2$$

極小時 $\frac{\partial(D^2)}{\partial m}=0$，$\frac{\partial(D^2)}{\partial c}=0$，可得

$$\sum_{i=1}^{n}\frac{x_iy_i}{\sigma_i^2}=c\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}+m\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\quad\cdots\cdots\cdots\cdots(1)$$

$$\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}=c\sum_{i=1}^{n}\frac{1}{\sigma_i^2}+m\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\quad\cdots\cdots\cdots\cdots(2)$$

由(1)(2)解聯立方程式可求斜率 m 和截距 c，y 與 x 之關係也就確立了，上兩式是通式，在精密實驗中，一個 y_i 值往往是數次量度下求得的平均值，因此 σ_i 並不完全相同，這時可用矩陣來解卽(1)、(2)式可寫爲 $Y=X\binom{c}{m}$ ∴ $\binom{c}{m}=X^{-1}Y$

這裡

$$Y=\begin{pmatrix}\sum_{i=1}^{n}\frac{x_iy_i}{\sigma_i^2}\\ \sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\end{pmatrix}$$

$$X=\begin{pmatrix}\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2} & \sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\\ \sum_{i=1}^{n}\frac{1}{\sigma_i^2} & \sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\end{pmatrix}$$

當然這是很繁的計算，尤其 n 相當大時，因此一般都是交由電腦處理，本文所附之電腦程式卽爲專爲處理此問題用的，它不但可處理 linear-linear，linear-log也可處理 log-log 的線性最小二次方式固定問題。

在一般情況下可假設 σ_i 均相等，如前面所述虎克定律之印證，可能每一點只測一次，因此令 σ_i 均相同是合理的，因此(1)，(2)可簡化成

$$\sum_{i=1}^{n}x_iy_i=c\sum_{i=1}^{n}x_i+m\sum_{i=1}^{n}x_i^2\qquad(3)$$

$$\sum_{i=1}^{n}y_i=cn+m\sum_{i=1}^{n}x_i\qquad(4)$$

這時計算斜率 m 及截距 c 就簡單多了，一般中學實驗教學如前述的虎克定律，如何將所獲之數據組繪出最佳之直線，就需從(3)、(4)方程式算出 m 及 c，就可正確的繪出我們所要的直線了。

至於多項式的最小二次方式固定法，非常繁複，幾乎無法用人工計算，非得用電腦處理不可，因中學實驗教學較少用到，就不擬介紹了。

參考資料

1. Brown, J. W., AV Instruction Media and Methods, McGraw-Hill, N.Y. 1969.
2. Bevington, P. R., Data Reduction & Errow Analysis for the Physics Sciences, Mc-Graw-Hill, N.Y. 1969.
3. Young, Hugh D., Statistical Treatment of Experimental Data, McGraw-Hill, N.Y. 1962.

附表：線性最小二次方式固定法電腦程式

```
:HELLO MAJOR.PHYSICS
USER PASSWORD?
MMMMMMMM
SESSION NUMBER = #S236
FRI, JUL 8, 1977, 3:03 PM
HP32002A.01.00

**** WELCOME TO WACC **** A SERVICE OF UW-RIVER FALLS ****
:BASIC

HP32101B.00.05(4WD) BASIC (C)HEWLETT-PACKARD CO 1976
>GET CH1
>LIST
CH1
  10 PRINT "THIS PROGRAM IS THE LINEAR LEAST SQUARES FITTING"


  20 REM THIS PROGRAM WAS WRITTEN BY CHING-SONG SHERN
  30 DIM P[2,1],Q[2,2],R[2,2],T[2,1]
  35 DIM U[2,1],V[2,2]
  40 DIM X[100],Y[100],S[100]
  41 DIM F[100],G[100],H[100]
  50 PRINT "WILL YOU DO LINEAR-LINEAR OR LINEAR-LN OR LN-LN FITTING?"
  51 PRINT "TYPE 1,2 OR 3 RESPECTIVELY"
  54 INPUT J
  60 PRINT "HOW MANY DATA PAIRS DO YOU HAVE?"
  65 INPUT N
  70 PRINT &
     "NOW INPUT YOUR DATA GROUPS.ORDER OF X,Y,STANDARD DEVIATION OF Y,"
  71 PRINT "SAME GROUP OF DATA WRITTEN IN SAME LINE"
  73 FOR I=1 TO N
  75   INPUT F[I],G[I],H[I]
  76 NEXT I
  80 PRINT LIN(2),&
     "    X            Y          S          LN X        LN Y        LN S
     "LIN(& 1)
  82 FOR I=1 TO N
  84   PRINT USING 85;F[I],G[I],H[I],LOG(F[I]),LOG(G[I]),LOG(H[I])
  85   IMAGE 6(4D,3D,2X)
  86 NEXT I
  90 IF J=1 THEN 100
  92 IF J=2 THEN 110
  94 IF J=3 THEN 120
  96 PRINT "BAD INPUT"
  98 GOTO 500
 100 FOR I=1 TO N
 102   X[I]=F[I]
 103   Y[I]=G[I]
 104   S[I]=H[I]
 106 NEXT I
 108 GOSUB 140
 109 GOTO 420
 110 FOR I=1 TO N
 112   X[I]=F[I]
 113   Y[I]=LOG(G[I])
 114   S[I]=LOG(H[I])
 116 NEXT I
```

```
118 GOSUB 140
119 GOTO 420
120 FOR I=1 TO N
122 X[I]=LOG(F[I])
123 Y[I]=LOG(G[I])
124 S[I]=LOG(H[I])
126 NEXT I
128 GOSUB 140
129 GOTO 420
140 LET A=B=C=D=E=0
142 FOR I=1 TO N
145   LET H=S[I]**2
150   LET A=Y[I]/H+A
160   LET B=Y[I]*X[I]/H+B
170   LET C=1/H+C
180   LET D=X[I]/H+D
190   LET E=X[I]**2/H+E
200 NEXT I
210 U[1,1]=A
220 U[2,1]=B
230 MAT P=U
240 V[1,1]=C
250 V[1,2]=V[2,1]=D
270 V[2,2]=E
280 MAT Q=V
290 MAT R=INV(Q)
300 MAT T=R*P
310 LET K=T[1,1]
320 LET M=T[2,1]
321 LET Y3=K
330 PRINT LIN(2),"Y=MX+K"
340 PRINT "M=",M
350 PRINT "K=";K
355 PRINT "THE INTERCEPT=",Y3,LIN(2)
360 PRINT &
    "CHOOSE TWO POINTS OF X. NOTICE:IF DO LN-LN FITTING,MUST INPUT"
361 PRINT "LN X VALUE."
370 INPUT X1,X2
380 LET Y1=M*X1+K
390 LET Y2=M*X2+K
400 PRINT LIN(1),"X1,Y1="X1,Y1
410 PRINT "X2,Y2=",X2,Y2
415 RETURN
420 PRINT LIN(2),"MORE DATA? TYPE 'YES'OR'NO'"
430 N$="NO"
440 Y$="YES"
450 INPUT A$
460 IF A$=Y$ THEN 50
470 IF A$=N$ THEN 490
480 PRINT "BAD INPUT, I NEED THE ANSWER 'YES' OR 'NO'"
485 GOTO 420
490 PRINT "GOOD BYE"
500 END
>EXIT

 END OF SUBSYSTEM

:BYE

CPU (SEC)=2
CONNECT (MIN)=3
FRI, JUL 8, 1977, 3:06 PM
END OF SESSION
```